Created by Neil Yetz & Gemma Wallace

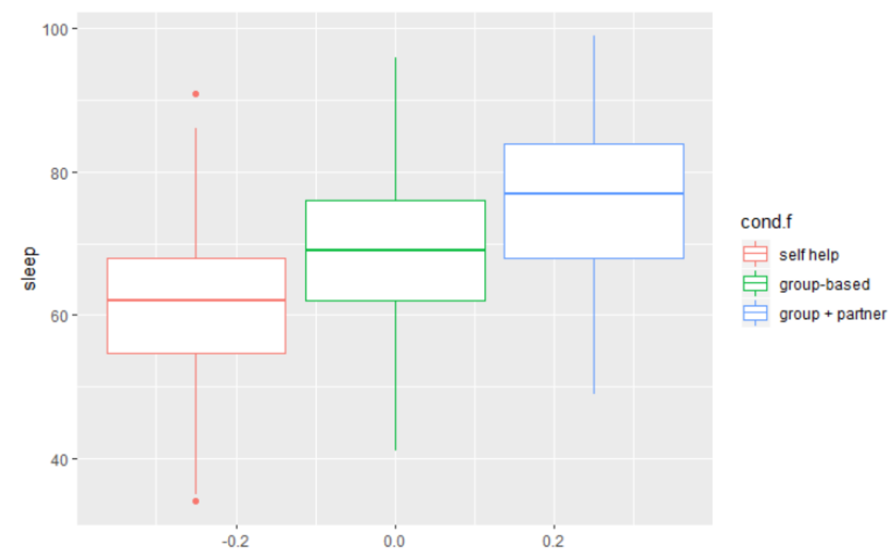
Answer Key

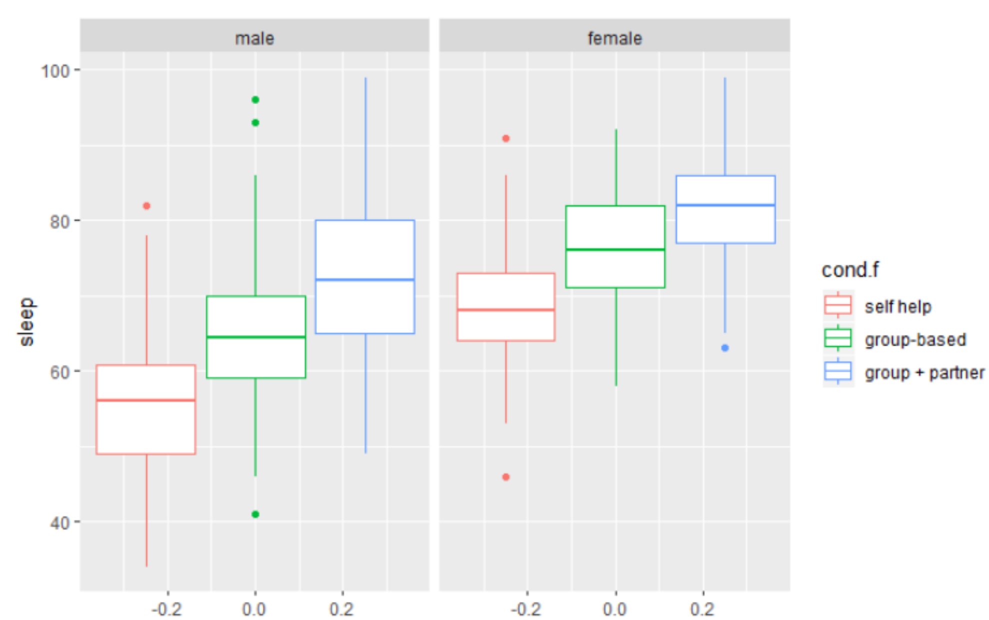
```
---
title: "R Notebook"
output:
  html_document:
    df_print: paged
---

# Load libraries
```{r, message = FALSE}
library(tidyverse)
library(psych)
library(car)
library(lsr)
library(MBESS)
```

Import Data
```{r, message = FALSE}

slp <- read_csv("slpdata.csv")

```


Factor grouping variables
```{r}
slp <- mutate(slp,
 female = ifelse(sex == 1, 0, 1),
 female.f = factor(female, levels = c(0,1), labels = c("male", "female")))

slp <- mutate(slp, cond.f = factor(cond, levels = c(1,2,3),
                                   labels = c("self help", "group-based", "group + partner")))
```


Calculate descriptives
```

```
# Calculate descriptives
# For whole dataset
```{r}
describe(slp)
```
```

	vars <dbl>	n <dbl>	mean <dbl>	sd <dbl>	median <dbl>	trimmed <dbl>	mad <dbl>	min <dbl>	max <dbl>
cond	1	600	2.00	0.82	2.00	2.00	1.48	1.00	3.00
prior	2	600	0.72	0.45	1.00	0.78	0.00	0.00	1.00
age	3	600	44.94	12.87	45.20	45.12	16.46	20.00	67.80
anxiety	4	600	3.88	0.90	3.86	3.89	0.93	1.05	6.84
hygiene	5	600	5.99	1.57	6.05	6.04	1.57	1.68	9.74
support	6	600	3.04	0.68	2.96	3.02	0.73	1.09	4.91
sleep	7	600	68.88	12.14	69.00	69.09	11.86	34.00	99.00
lifesat	8	600	4.06	0.92	4.05	4.04	0.96	1.68	6.61
sex	9	600	1.41	0.49	1.00	1.39	0.00	1.00	2.00
id	10	600	300.50	173.35	300.50	300.50	222.39	1.00	600.00

1-10 of 13 rows | 1-10 of 13 columns   Previous 1 2 Next

```
## Summarize descriptives grouping variables
```{r}
aggregate(x=slp$sleep, by=list(slp$female.f, slp$cond.f), FUN=mean)
```
```

Group.1 <fctr>	Group.2 <fctr>	x <dbl>
male	self help	54.86792
female	self help	68.34043
male	group-based	65.01538
female	group-based	76.70000
male	group + partner	72.75214
female	group + partner	81.36145

6 rows

```
# Visualize the data
```

````
# Visualize the data
## Create boxplots of sleep efficiency across treatment groups
```{r}
ggplot(slp, aes(y = sleep, color = cond.f)) +
  geom_boxplot()
```
````




````
## Create boxplots of sleep efficiency across treatment groups and sex
```{r}
ggplot(slp, aes(y = sleep, color = cond.f)) +
  geom_boxplot() +
  facet_wrap(~female.f)
```
````



The faceted boxplot tells us that the group + partner condition experiences the highest scores on sleep for both genders. However, females seem to have higher sleep scores as compared to males.

```
# Conduct a factorial ANOVA
```{r}
model <- lm(sleep ~ female.f + cond.f + female.f*cond.f, data = slp)
Anova(model, type = 3)
```
```

```
Anova Table (Type III tests)

Response: sleep
                Sum Sq  Df   F value  Pr(>F)    
(Intercept)     319112   1 4073.6899 < 2e-16 ***
female.f          9043   1  115.4370 < 2e-16 ***
cond.f           17836   2  113.8443 < 2e-16 ***
female.f:cond.f    593   2    3.7865 0.02322 *  
Residuals        46531 594                      
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
## Calculate effect sizes
```{r}

etaSquared(model, type = 3, anova = FALSE)
```
```

```
                     eta.sq eta.sq.part
female.f        0.102436146  0.16271641
cond.f          0.202045480  0.27709833
female.f:cond.f 0.006720116  0.01258869
```

b.) In the white space, interpret the η^2 and partial η^2 values for each main effect and the interaction effect.

The η^2 is the effect of the variable on the entire variance of the outcome. Therefore, female, explains 10% of the variance in sleep, condition explains 20% of the variance in sleep and the interaction of the two explains <1% of the variance in sleep.

The partial η^2 is the effect of the variable on the outcome *after* partialling out the variance from the other variables in our model. Therefore, female explains 16% of the partialled out variance in sleep, condition explains 28% of the partialled variance in sleep and the interaction of female and condition explains 1% of the partialled variance in sleep.

Created by Neil Yetz & Gemma Wallace

```{r}


ci.pvaf(F.value = 115.4370, df.1 =1, df.2 = 594, N = 600, conf.level = .95)
ci.pvaf(F.value = 113.8443, df.1 =2, df.2 = 594, N = 600, conf.level = .95)
ci.pvaf(F.value = 3.7865,    df.1 =2, df.2 = 594, N = 600, conf.level = .95)


```

```
$Lower.Limit.Proportion.of.Variance.Accounted.for
[1] 0.1117234

$Probability.Less.Lower.Limit
[1] 0.025

$Upper.Limit.Proportion.of.Variance.Accounted.for
[1] 0.2143099

$Probability.Greater.Upper.Limit
[1] 0.025

$Actual.Coverage
[1] 0.95

$Lower.Limit.Proportion.of.Variance.Accounted.for
[1] 0.2173461

$Probability.Less.Lower.Limit
[1] 0.025

$Upper.Limit.Proportion.of.Variance.Accounted.for
[1] 0.330051

$Probability.Greater.Upper.Limit
[1] 0.025

$Actual.Coverage
[1] 0.95

$Lower.Limit.Proportion.of.Variance.Accounted.for
[1] 6.757303e-05

$Probability.Less.Lower.Limit
[1] 0.025

$Upper.Limit.Proportion.of.Variance.Accounted.for
[1] 0.03371934

$Probability.Greater.Upper.Limit
[1] 0.025

$Actual.Coverage
[1] 0.95
```

Each of the eta squared confidence intervals tell us the range of plausible values of the partial eta squared. If we ran this experiment 100 times, 95% of our results would likely fall within the range of these values.

Female: Partial η^2 = .16, 95% CI = [.11, .21]
Condition: Partial η^2 = .28, 95% CI = [.22, .33]
Interaction: Partial η^2 = .01, 95% CI = [<.01, .03]